

AFFECTIVE TELEVISION MONITORING AND CONTROL

5

TECHNICAL FIELD OF THE INVENTION

The present invention is directed, in general, to television audience feedback collection systems and, more specifically, to a system and method for using human-response inputs to automatically create a viewer profile.

10

BACKGROUND OF THE INVENTION

Modern television systems effectively process an amazing quantity of audio-visual input to present viewers with high quality programs from a wide variety of sources. Many of these sources are broadcast networks that send programming through cables or over the airwaves. Other sources may be more local or limited. Private networks, for example, may transmit programming that is intended only for a limited audience, and many homes are equipped with videocassette recorders (VCRs) or other recording devices which, once they hold recorded material, are programming sources themselves.

Modern viewers, therefore, have a great deal of programming choice. As most of the programming is generated for commercial

reasons, program producers have an interest in determining what viewers want or like to watch so that popular programming can be used as a guide to designing programs for the future. In addition, in order to cope with the huge quantity of available programming 5 there are various devices for allowing viewers to selectively focus on a relatively small portion of it, should they choose to do so. For example, a viewer may simply choose to subscribe to only certain channels of programming. For whatever purpose it is collected, however, it is often useful to collect data regarding 10 the viewing preferences of a television user (viewer). There are several ways in which to collect helpful data on what it is that people prefer to watch. First of all, viewers can simply be asked. A given program might be shown to a test audience, and then its members queried to determine their reaction. Although generally efficacious, this approach does have several drawbacks. For one 15 thing, it requires a great many interviewers to ask the questions and record the answers. Then the answers must be properly interpreted if an accurate reading of viewer likes and dislikes is to be made. For these reasons, this approach, in general, is of value only with a limited test audience. Because the viewer 20 response is to be collected from only this limited audience, these

viewers must be properly selected to ensure their responses will be representative of the viewing public at large.

Interview personnel are not required, of course, if the test viewers are simply asked to supply answers to predetermined written 5 questions, write a description of their response, or "vote" for which of a plurality of programs they enjoyed the most. This approach is also subject, however, to even more errors in interpretation and carries with it a problem that might be referred to as respondent fatigue. This is the situation where the test viewer may at some point get tired of answering questions or filling out forms, and, if so, complete them carelessly because all they now desire to do is fulfill the agreed assignment. There also exists the problem of purposely misdirected answers, where the viewer senses the purpose of a question and, for some reason, provides a misleading response. For example, programming that is 15 attractive to a viewer might be explicit or violent enough that the viewer does not want to admit its pleasing nature. However well motivated, such behavior corrupts the testing process.

And, of course, the results are only as good as the questions 20 asked and the answers given. Any interview script or written questionnaire must be carefully designed to yield accurate

responses, even from a sincere and earnest respondent. All of this requires time, money and other resources, and so only a limited audience may be tested. However carefully the testers try to assemble test audiences to provide statistically valid samples,
5 they are limited to those having the willingness to respond to questions. This problem exists even where viewers are called at home and questioned about how they watch, and about their viewing habits in general. And with any of the methods described above, the problem persists that generally the viewer bases the report of
10 their reaction on the viewed program as a whole, and not on the specific segments of the program that may have produced a positive or negative response. In many cases, such responses may be intuitive or subliminal, and even an earnest viewer may not be able
15 to pinpoint exactly what portion of a program was most desirable. Finally, the questions and answers, or other inquiry mechanism must either follow the entire program or interrupt it before it is finished. In either case, the integrity of assessing the audience reaction is compromised. As mentioned above, viewer feedback may also be used merely to simplify their own channel selection
20 process. But even where viewers simply try to remember which programs they like and adjust their channel selector, the results

may be far from perfect. Their recollection may be faulty, and the channels they select may not be well matched to their real preferences. In any case, they must perform the selection manually, and may procrastinate in doing so because it is too much
5 trouble.

Needed then is an unintrusive and automatic method of gauging audience reaction to television and similar programming that can be implemented over a wide testing audience, and that provides accurate feedback to the party conducting the test, or simply to an automatic program-selection assistance system. The system and method of the present invention provide just such a solution.

SUMMARY OF THE INVENTION

To address the above-discussed deficiencies of the prior art,
15 it is a primary object of the present invention to provide, for use with a television system, a system and method for collecting and analyzing, accurately and unobtrusively, audience reaction to specific programs and program segments. One or more viewer-feedback sensors are placed in proximity to each test viewer to
20 collect response information. The response data is translated into digital form for storage and analysis and stored at a system

database. The response information stored on the database is time-stamped so that it can be corrected with a specific program segment. A set top box is coupled to one or more programming sources. A control program either selects the programming for the
5 viewers to watch or allows them to make their own selection.

Once a sufficient quantity of material has been viewed, the data associated with an identifiable program segment is assessed. A reporting device delivers the results of the analysis.

In a particularly preferred embodiment, the reported results
10 are used to inform the testing process so that viewer responses can be confirmed.

In another particularly preferred embodiment, the viewer responses are associated with a viewer preference level (indicative
15 of viewer likes or dislikes, which are inferred from the collected response information).

Additional objects and advantages of the present invention will be more fully described in the DETAILED DESCRIPTION of the invention.

The foregoing has outlined rather broadly the features and
20 technical advantages of the present invention so that those skilled in the art may better understand the detailed description of the

invention that follows. Additional features and advantages of the invention will be described hereinafter that form the subject of the claims of the invention. Those skilled in the art should appreciate that they may readily use the conception and the 5 specific embodiment disclosed as a basis for modifying or designing other structures for carrying out the same purposes of the present invention. Those skilled in the art should also realize that such equivalent constructions do not depart from the spirit and scope of the invention in its broadest form.

10 Before undertaking the DETAILED DESCRIPTION, it may be advantageous to set forth definitions of certain words and phrases used throughout this patent document: the terms "include" and "comprise," as well as derivatives thereof, mean inclusion without limitation; the term "or," is inclusive, meaning and/or; the 15 phrases "associated with" and "associated therewith," as well as derivatives thereof, may mean to include, be included within, interconnect with, contain, be contained within, connect to or with, couple to or with, be communicable with, cooperate with, interleave, juxtapose, be proximate to, be bound to or with, have, 20 have a property of, or the like; and the term "controller" means any device, system or part thereof that controls at least one

operation, such a device may be implemented in hardware, firmware or software, or some combination of at least two of the same. It should be noted that the functionality associated with any particular controller may be centralized or distributed, whether 5 locally or remotely. Definitions for certain words and phrases are provided throughout this patent document, those of ordinary skill in the art should understand that in many, if not most instances, such definitions apply to prior, as well as future uses of such defined words and phrases.

10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
260
265
270
275
280
285
290
295
300
305
310
315
320
325
330
335
340
345
350
355
360
365
370
375
380
385
390
395
400
405
410
415
420
425
430
435
440
445
450
455
460
465
470
475
480
485
490
495
500
505
510
515
520
525
530
535
540
545
550
555
560
565
570
575
580
585
590
595
600
605
610
615
620
625
630
635
640
645
650
655
660
665
670
675
680
685
690
695
700
705
710
715
720
725
730
735
740
745
750
755
760
765
770
775
780
785
790
795
800
805
810
815
820
825
830
835
840
845
850
855
860
865
870
875
880
885
890
895
900
905
910
915
920
925
930
935
940
945
950
955
960
965
970
975
980
985
990
995
1000
1005
1010
1015
1020
1025
1030
1035
1040
1045
1050
1055
1060
1065
1070
1075
1080
1085
1090
1095
1100
1105
1110
1115
1120
1125
1130
1135
1140
1145
1150
1155
1160
1165
1170
1175
1180
1185
1190
1195
1200
1205
1210
1215
1220
1225
1230
1235
1240
1245
1250
1255
1260
1265
1270
1275
1280
1285
1290
1295
1300
1305
1310
1315
1320
1325
1330
1335
1340
1345
1350
1355
1360
1365
1370
1375
1380
1385
1390
1395
1400
1405
1410
1415
1420
1425
1430
1435
1440
1445
1450
1455
1460
1465
1470
1475
1480
1485
1490
1495
1500
1505
1510
1515
1520
1525
1530
1535
1540
1545
1550
1555
1560
1565
1570
1575
1580
1585
1590
1595
1600
1605
1610
1615
1620
1625
1630
1635
1640
1645
1650
1655
1660
1665
1670
1675
1680
1685
1690
1695
1700
1705
1710
1715
1720
1725
1730
1735
1740
1745
1750
1755
1760
1765
1770
1775
1780
1785
1790
1795
1800
1805
1810
1815
1820
1825
1830
1835
1840
1845
1850
1855
1860
1865
1870
1875
1880
1885
1890
1895
1900
1905
1910
1915
1920
1925
1930
1935
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
2025
2030
2035
2040
2045
2050
2055
2060
2065
2070
2075
2080
2085
2090
2095
2100
2105
2110
2115
2120
2125
2130
2135
2140
2145
2150
2155
2160
2165
2170
2175
2180
2185
2190
2195
2200
2205
2210
2215
2220
2225
2230
2235
2240
2245
2250
2255
2260
2265
2270
2275
2280
2285
2290
2295
2300
2305
2310
2315
2320
2325
2330
2335
2340
2345
2350
2355
2360
2365
2370
2375
2380
2385
2390
2395
2400
2405
2410
2415
2420
2425
2430
2435
2440
2445
2450
2455
2460
2465
2470
2475
2480
2485
2490
2495
2500
2505
2510
2515
2520
2525
2530
2535
2540
2545
2550
2555
2560
2565
2570
2575
2580
2585
2590
2595
2600
2605
2610
2615
2620
2625
2630
2635
2640
2645
2650
2655
2660
2665
2670
2675
2680
2685
2690
2695
2700
2705
2710
2715
2720
2725
2730
2735
2740
2745
2750
2755
2760
2765
2770
2775
2780
2785
2790
2795
2800
2805
2810
2815
2820
2825
2830
2835
2840
2845
2850
2855
2860
2865
2870
2875
2880
2885
2890
2895
2900
2905
2910
2915
2920
2925
2930
2935
2940
2945
2950
2955
2960
2965
2970
2975
2980
2985
2990
2995
3000
3005
3010
3015
3020
3025
3030
3035
3040
3045
3050
3055
3060
3065
3070
3075
3080
3085
3090
3095
3100
3105
3110
3115
3120
3125
3130
3135
3140
3145
3150
3155
3160
3165
3170
3175
3180
3185
3190
3195
3200
3205
3210
3215
3220
3225
3230
3235
3240
3245
3250
3255
3260
3265
3270
3275
3280
3285
3290
3295
3300
3305
3310
3315
3320
3325
3330
3335
3340
3345
3350
3355
3360
3365
3370
3375
3380
3385
3390
3395
3400
3405
3410
3415
3420
3425
3430
3435
3440
3445
3450
3455
3460
3465
3470
3475
3480
3485
3490
3495
3500
3505
3510
3515
3520
3525
3530
3535
3540
3545
3550
3555
3560
3565
3570
3575
3580
3585
3590
3595
3600
3605
3610
3615
3620
3625
3630
3635
3640
3645
3650
3655
3660
3665
3670
3675
3680
3685
3690
3695
3700
3705
3710
3715
3720
3725
3730
3735
3740
3745
3750
3755
3760
3765
3770
3775
3780
3785
3790
3795
3800
3805
3810
3815
3820
3825
3830
3835
3840
3845
3850
3855
3860
3865
3870
3875
3880
3885
3890
3895
3900
3905
3910
3915
3920
3925
3930
3935
3940
3945
3950
3955
3960
3965
3970
3975
3980
3985
3990
3995
4000
4005
4010
4015
4020
4025
4030
4035
4040
4045
4050
4055
4060
4065
4070
4075
4080
4085
4090
4095
4100
4105
4110
4115
4120
4125
4130
4135
4140
4145
4150
4155
4160
4165
4170
4175
4180
4185
4190
4195
4200
4205
4210
4215
4220
4225
4230
4235
4240
4245
4250
4255
4260
4265
4270
4275
4280
4285
4290
4295
4300
4305
4310
4315
4320
4325
4330
4335
4340
4345
4350
4355
4360
4365
4370
4375
4380
4385
4390
4395
4400
4405
4410
4415
4420
4425
4430
4435
4440
4445
4450
4455
4460
4465
4470
4475
4480
4485
4490
4495
4500
4505
4510
4515
4520
4525
4530
4535
4540
4545
4550
4555
4560
4565
4570
4575
4580
4585
4590
4595
4600
4605
4610
4615
4620
4625
4630
4635
4640
4645
4650
4655
4660
4665
4670
4675
4680
4685
4690
4695
4700
4705
4710
4715
4720
4725
4730
4735
4740
4745
4750
4755
4760
4765
4770
4775
4780
4785
4790
4795
4800
4805
4810
4815
4820
4825
4830
4835
4840
4845
4850
4855
4860
4865
4870
4875
4880
4885
4890
4895
4900
4905
4910
4915
4920
4925
4930
4935
4940
4945
4950
4955
4960
4965
4970
4975
4980
4985
4990
4995
5000
5005
5010
5015
5020
5025
5030
5035
5040
5045
5050
5055
5060
5065
5070
5075
5080
5085
5090
5095
5100
5105
5110
5115
5120
5125
5130
5135
5140
5145
5150
5155
5160
5165
5170
5175
5180
5185
5190
5195
5200
5205
5210
5215
5220
5225
5230
5235
5240
5245
5250
5255
5260
5265
5270
5275
5280
5285
5290
5295
5300
5305
5310
5315
5320
5325
5330
5335
5340
5345
5350
5355
5360
5365
5370
5375
5380
5385
5390
5395
5400
5405
5410
5415
5420
5425
5430
5435
5440
5445
5450
5455
5460
5465
5470
5475
5480
5485
5490
5495
5500
5505
5510
5515
5520
5525
5530
5535
5540
5545
5550
5555
5560
5565
5570
5575
5580
5585
5590
5595
5600
5605
5610
5615
5620
5625
5630
5635
5640
5645
5650
5655
5660
5665
5670
5675
5680
5685
5690
5695
5700
5705
5710
5715
5720
5725
5730
5735
5740
5745
5750
5755
5760
5765
5770
5775
5780
5785
5790
5795
5800
5805
5810
5815
5820
5825
5830
5835
5840
5845
5850
5855
5860
5865
5870
5875
5880
5885
5890
5895
5900
5905
5910
5915
5920
5925
5930
5935
5940
5945
5950
5955
5960
5965
5970
5975
5980
5985
5990
5995
6000
6005
6010
6015
6020
6025
6030
6035
6040
6045
6050
6055
6060
6065
6070
6075
6080
6085
6090
6095
6100
6105
6110
6115
6120
6125
6130
6135
6140
6145
6150
6155
6160
6165
6170
6175
6180
6185
6190
6195
6200
6205
6210
6215
6220
6225
6230
6235
6240
6245
6250
6255
6260
6265
6270
6275
6280
6285
6290
6295
6300
6305
6310
6315
6320
6325
6330
6335
6340
6345
6350
6355
6360
6365
6370
6375
6380
6385
6390
6395
6400
6405
6410
6415
6420
6425
6430
6435
6440
6445
6450
6455
6460
6465
6470
6475
6480
6485
6490
6495
6500
6505
6510
6515
6520
6525
6530
6535
6540
6545
6550
6555
6560
6565
6570
6575
6580
6585
6590
6595
6600
6605
6610
6615
6620
6625
6630
6635
6640
6645
6650
6655
6660
6665
6670
6675
6680
6685
6690
6695
6700
6705
6710
6715
6720
6725
6730
6735
6740
6745
6750
6755
6760
6765
6770
6775
6780
6785
6790
6795
6800
6805
6810
6815
6820
6825
6830
6835
6840
6845
6850
6855
6860
6865
6870
6875
6880
6885
6890
6895
6900
6905
6910
6915
6920
6925
6930
6935
6940
6945
6950
6955
6960
6965
6970
6975
6980
6985
6990
6995
7000
7005
7010
7015
7020
7025
7030
7035
7040
7045
7050
7055
7060
7065
7070
7075
7080
7085
7090
7095
7100
7105
7110
7115
7120
7125
7130
7135
7140
7145
7150
7155
7160
7165
7170
7175
7180
7185
7190
7195
7200
7205
7210
7215
7220
7225
7230
7235
7240
7245
7250
7255
7260
7265
7270
7275
7280
7285
7290
7295
7300
7305
7310
7315
7320
7325
7330
7335
7340
7345
7350
7355
7360
7365
7370
7375
7380
7385
7390
7395
7400
7405
7410
7415
7420
7425
7430
7435
7440
7445
7450
7455
7460
7465
7470
7475
7480
7485
7490
7495
7500
7505
7510
7515
7520
7525
7530
7535
7540
7545
7550
7555
7560
7565
7570
7575
7580
7585
7590
7595
7600
7605
7610
7615
7620
7625
7630
7635
7640
7645
7650
7655
7660
7665
7670
7675
7680
7685
7690
7695
7700
7705
7710
7715
7720
7725
7730
7735
7740
7745
7750
7755
7760
7765
7770
7775
7780
7785
7790
7795
7800
7805
7810
7815
7820
7825
7830
7835
7840
7845
7850
7855
7860
7865
7870
7875
7880
7885
7890
7895
7900
7905
7910
7915
7920
7925
7930
7935
7940
7945
7950
7955
7960
7965
7970
7975
7980
7985
7990
7995
8000
8005
8010
8015
8020
8025
8030
8035
8040
8045
8050
8055
8060
8065
8070
8075
8080
8085
8090
8095
8100
8105
8110
8115
8120
8125
8130
8135
8140
8145
8150
8155
8160
8165
8170
8175
8180
8185
8190
8195
8200
8205
8210
8215
8220
8225
8230
8235
8240
8245
8250
8255
8260
8265
8270
8275
8280
8285
8290
8295
8300
8305
8310
8315
8320
8325
8330
8335
8340
8345
8350
8355
8360
8365
8370
8375
8380
8385
8390
8395
8400
8405
8410
8415
8420
8425
8430
8435
8440
8445
8450
8455
8460
8465
8470
8475
8480
8485
8490
8495
8500
8505
8510
8515
8520
8525
8530
8535
8540
8545
8550
8555
8560
8565
8570
8575
8580
8585
8590
8595
8600
8605
8610
8615
8620
8625
8630
8635
8640
8645
8650
8655
8660
8665
8670
8675
8680
8685
8690
8695
8700
8705
8710
8715
8720
8725
8730
8735
8740
8745
8750
8755
8760
8765
8770
8775
8780
8785
8790
8795
8800
8805
8810
8815
8820
8825
8830
8835
8840
8845
8850
8855
8860
8865
8870
8875
8880
8885
8890
8895
8900
8905
8910
8915
8920
8925
8930
8935
8940
8945
8950
8955
8960
8965
8970
8975
8980
8985
8990
8995
9000
9005
9010
9015
9020
9025
9030
9035
9040
9045
9050
9055
9060
9065
9070
9075
9080
9085
9090
9095
9100
9105
9110
9115
9120
9125
9130
9135
9140
9145
9150
9155
9160
9165
9170
9175
9180
9185
9190
9195
9200
9205
9210
9215
9220
9225
9230
9235
9240
9245
9250
9255
9260
9265
9270
9275
9280
9285
9290
9295
9300
9305
9310
9315
9320
9325
9330
9335
9340
9345
9350
9355
9360
9365
9370
9375
9380
9385
9390
9395
9400
9405
9410
9415
9420
9425
9430
9435
9440
9445
9450
9455
9460
9465
9470
9475
9480
9485
9490
9495
9500
9505
9510
9515
9520
952

BRIEF DESCRIPTION OF THE DRAWINGS

For a more complete understanding of the present invention, and the advantages thereof, reference is now made to the following descriptions taken in conjunction with the accompanying drawings, 5 wherein like numbers designate like objects, and in which:

FIGURE 1 illustrates an exemplary television-receiver system, according to one embodiment of the present invention;

FIGURE 2 illustrates an exemplary system for collecting viewer responses to program segments that is implemented in a system analogous to the exemplary television receiver system of FIGURE 1, according to a similar embodiment of the present invention;

FIGURE 3 is a block diagram illustrating a video processor that may be used to process input from a video camera according to a preferred embodiment of the present invention;

FIGURE 4 is a system diagram illustrating the interaction of sensory inputs according to one embodiment of the present invention, as well as possible applications for using the gathered and processed sensory data; and

FIGURE 5 is a flow chart illustrating a method of collecting, 20 processing, and using sensory-feedback information according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

FIGURES 1 through 5, discussed below, and the various embodiments used to describe the principles of the present invention in this patent document are by way of illustration only and should not be construed in any way to limit the scope of the invention. In the description of the exemplary embodiment that follows, the present invention is integrated into, or is used in connection with, a digital television receiver. However, this embodiment is by way of example only and should not be construed to limit the scope of the present invention to digital television receivers. In fact, those skilled in the art will recognize that the exemplary embodiment of the present invention may easily be modified for use in analog television receivers and other similar devices as well.

Note that television "program" or "programming" will be used generally herein to refer to programs, or portions thereof, that may be displayed on a television receiver or similar device including, for example, broadcast or cable TV, prerecorded video tape or DVD presentations, and streaming multimedia sent over a telecommunications or computer network. As used herein, "television programming" also includes audio programming and

textual or multimedia presentations, whether they accompany an actual television broadcast or not. For example, radio broadcasts and multimedia web page presentations are included as well (and in appropriate instances "view" or "viewer" will refer to simply 5 listening, or listening and reading as well as actually seeing video images).

FIGURE 1 illustrates exemplary television system 100, according to one embodiment of the present invention. Exemplary television system 100 comprises television receiver 105, set top box 150 with infrared (IR) detector 160, and video-recording device 140 (which, although typically so-called, usually records program audio, as well). As will be explained below in greater detail, the present invention provides a system for collecting and using viewer-response as feedback for research purposes and for the enhanced operation of television receiver 105, especially with regard to the programming material selected for display and for recording.

Television receiver 105 comprises display screen 110 for displaying television programs provided by a cable or satellite 20 television program service provider, and for displaying previously recorded material. Receiver 105 also includes infrared (IR)

sensor 115, and a set of manual controls 120, as indicated by a surrounding dotted line. Manual controls 120 may include, without limitation, a power button, a volume control button, vertical and horizontal control buttons and one or more channel selection 5 buttons. Infrared (IR) sensor 115 receives IR control signals from an optional hand-held remote control 125 that may be operated by the television viewer. Typically, IR control signals from remote control 125 that are detected by IR sensor 115 are processed within television receiver 105 in order to change the channel being viewed 10 on display screen 110, to increase or to decrease the volume, to turn television receiver 105 on and off, and the like. Optionally, the IR control signals detected by IR sensor 115 may be relayed to set top box 150 in order to control the operation of set top 15 box 150.

Set top box 150 performs conventional tuning and demodulation 20 of incoming signals from a cable or satellite television program service provider to produce, for example, a stream of Moving Picture Experts Group (MPEG) encoded digital data from which video signals may be derived. Alternatively, television receiver 105 may directly receive an incoming television broadcast signal from an external antenna (not shown). In this alternative embodiment of

the invention, television receiver 105 performs conventional tuning and demodulation of incoming RF signals received from the antenna to produce, for example, a stream of MPEG encoded digital data from which video signals may be derived. The external antenna, if 5 present, may also be connected to set top box 150.

Set top box 150 also includes infrared (IR) sensor 160. IR sensor 160 receives infrared (IR) control signals from hand-held remote control 125 operated by the television viewer. Preferably, remote control 125 that controls set top box 150 is the same remote control 125 that operates television receiver 105. Typically, 10 IR control signals that are detected by IR sensor 160 are processed within set top box 150 in order to change the channel being transmitted to television receiver 105 for viewing on display screen 110, to turn set top box 150 (and/or television receiver 105) on and off, and to adjust other television controls. 15

Video recorder 140 is a device that records programs for the viewer to watch at another time. It may be connected to television receiver 105 or, preferably, directly to set top box 150, and is capable of recording what is being displayed on display 110, but 20 can also record when the display is off. A particular advantage is obtained when television system 100 is configured such that one

program may be recorded while a second is being displayed. Video recorder 140 may be, but is not necessarily a video cassette recorder (VCR), and may be able to play back programs not received from set top box 150 or receiver 105, for example, using a 5 removable media such as a videotape or floppy disk, or may be connected directly or indirectly to a network that allows downloading of programs for recording and playback. In addition, recorder 140 might be coupled with the TV 105 or with the set-top box 150.

10 In an advantageous embodiment of the present invention, set top box 150 may be integrated into television receiver 105. The integration of set top box 150 and television receiver 105 commonly occurs, for example, where television receiver 105 is a 15 high-definition digital television (HDTV) receiver. Because the functions and principles of operation of set top box 150 are well known, television manufacturers frequently find it advantageous to integrate set top box 150 into television receiver 105, thereby reducing the amount of equipment, wiring, and set-up work required to be performed by the television viewer.

20 Preferably, television system 100 also includes a personal computer (PC) 170, which is in communication with both television

receiver 105, and set top box 150. PC 170 comprises central processing unit (CPU) 180, video monitor 190, and removable storage medium 195. PC 170 also comprises conventional elements (not shown) such as a keyboard, a mouse, an internal hard disk drive, and a 5 random access memory (RAM). In one embodiment of the present invention that uses PC 170, the system for collecting viewer feedback comprises CPU 180 and a program stored in the random access memory (RAM) (not shown) of CPU 180 or stored in the internal hard disk drive (not shown) of CPU 180. The program may also be stored on removable storage medium 195, which may be, for example, a 3.5 inch floppy diskette, a compact disk read only memory (CD ROM), a digital video disk (DVD), or a similar storage medium.

In addition, the viewer-feedback collection system of the 15 present invention comprises a plurality of viewer-response sensors, enumerated in FIGURE 1 as 132, 134, 136, and 138, although there may be any number of them. Sensor-signal receiver 185 receives the response signals produced by the viewer-response monitors 132-138 in their unprocessed state and processes them for storage and 20 analysis by PC 170.

In an advantageous embodiment of the present invention, the

system for collecting viewer responses may be an embedded system integrated into television receiver 105. In an alternate advantageous embodiment of the present invention, the system may be implemented in a set top box 150. In other words, the components 5 of television system 100 that are in FIGURE 1 shown separately, namely, sensor-signal receiver 185, PC 170, and set top box 150, may be incorporated into a single physical unit - such as receiver 105 - or their functions distributed in some other fashion.

FIGURE 2 illustrates an exemplary system for collecting viewer reactions to programs and program segments that is implemented in a system analogous to the exemplary television receiver system 100 of FIGURE 1, according to a similar embodiment of the present invention. Television receiver system 200 includes television receiver 205 having display 210. Set top box 250 receives programming information, as previously described, and transmits it to receiver 205. Note that television receiver system 200 is for clarity somewhat simplified from the embodiment illustrated in FIGURE 1, though the various components shown there may be present as well in system 200. Ideally, the viewer is seated in chair 215, 10 at least at the beginning of a test-audience section. By "test audience", it is meant that, as according to one embodiment of the 15

present invention, the system may be used in a controlled test situation. In an alternate embodiment, where the system is employed in a viewer's home, a similar chair 215 could nevertheless be provided. Note that the system and method of the present 5 invention is applicable for use in either application, and features described herein applicable to one will be deemed applicable to the other unless the need for distinction is either explicitly pointed out or readily apparent from the context.

Chair 215 is optionally equipped with sensors (not shown) for recording such aspects of the viewer's condition as temperature, posture, propensity to move around or to be still, and so on. A microphone (not shown) may also be present in or attached to chair 215 for recording responsive sounds such as laughter or conversational speech. Signals generated by these sensors are collected and converted into radio signals for wireless transmission to sensor-signal receiver 285. Chair 215 and sensor-signal receiver 285 are equipped with antennae 217 and 219, respectively, to aid radio communication between them. Wireless transmission is not required, of course, and sensor-signals may 15 also be transmitted by other means such as infrared or simply through a cable. Sensor shirt 220 provides another device for

collecting viewer feedback via involuntary or spontaneous responses. Note that the "involuntary" or "spontaneous" responses sought here are those that occur naturally for the viewer while watching the displayed programming. That is, these terms refer 5 generally to responses or reactions to the programming itself, at or near the time it is being displayed, and not to responses to a later query such as an interviewer's questions. A key advantage of the present invention, therefore, is the ability to receive reactions that can be connected directly with program segments, and 10 at the same time allow the program to continue uninterrupted. Sensor shirt 220 may be used in conjunction with or as an alternative to chair 215. Sensor shirt 220 contains one or more sensors to measure viewer conditions such as a body temperature sensor, heart rate monitor, perspiration detectors and so forth. 15 In a particularly preferred embodiment, sensor shirt 220 includes a galvactivator (not shown), which measures skin conductivity response (also known as the electrodermal response), taking advantage of the phenomenon that the skin momentarily becomes a better conductor of electricity when either external or internal 20 stimuli occur that are physiologically arousing. This phenomenon is more fully explained in Rosalind W. Picard & Jocelyn Scheirer,

The Galvactivator: A Glove that Senses and Communicates Skin Conductivity, PROCEEDINGS FROM THE 9TH INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION, NEW ORLEANS (August 2001), which is incorporated herein by reference. As should be apparent to one of ordinary skill in the art, such a sensor may also be incorporated into a glove or other article that is placed in contact with the viewer. This listing is intended to be illustrative and not limiting. Sensor shirt 220, glove, or other sensory device preferably includes wireless transmitter 222 for transmitting sensor data to sensor-signal receiver 285.

Other viewer-condition sensors present in the embodiment of FIGURE 2 include motion sensor 230 for sensing motion about (or even in-and-out of) the viewing room 201. There may be more than one motion sensor depending on the (field) of motion to be covered. Motion sensor 230 is ordinarily connected with sensor signal receiver 285 via a cable connection, but any of the other methods, such as those previously mentioned, may also be used. Video camera 240 is positioned to capture an image of a viewer seated in chair 215, and in particular, the head and face region of the viewer. It may be stationary or movable. In the latter case, a drive-control system (not shown) may be used to assist the video camera in

tracking and staying focused on the viewer.

Video camera 240 may be used to provide sensory information in a variety of ways. Unlike most of the other sensors previously recited (with the exception of the microphone), its signal output 5 can merely be saved onto a video recording device such as video recorder 140 shown in FIGURE 1. Normally, however, a separate video recorder (not shown) will be used if the viewer-reaction video images are to be recorded as such. In a particularly preferred embodiment, however, the video information is separately 10 processed to enhance its feedback value and to eliminate (or corroborate) any analysis of the video image performed by a human operator.

In a preferred embodiment, sensory information gathered through a plurality of sensors, such as the various devices 15 described, will be analyzed in combination to yield a more accurate profile of viewer response than would be obtained from simply looking at each sensory input in isolation. For example, a particular galvactic response may include arousal, but leave open 20 to question whether the elicited but unspecified emotion is good or bad. Used in combination with visual cues, however, valence can also be gauged; a smile, for instance, indicates this particular

state of arousal is a happy one. As another example, either a visual response such as a furrowed brow, or an audio response such as a question being asked (as determined by the speaker's rising end-of-sentence inflection), might indicate confusion or lack of 5 understanding. When the inputs are considered together, however, the likelihood increases that this is the correct determination.

FIGURE 3 is a block diagram illustrating a video processor 300 that may be used to process input from a video camera 240 according to a preferred embodiment of the present invention. Video processor 300 receives video input from camera 240 at input port 305. Preferably, video processor 300 also includes video recording medium 310 for recording the image captured by camera 240. Again, "video recording" is used herein for convenience. Typically, video recorders record audio and even perhaps accompanying text (such as closed-captioning) as well. In the illustrated embodiment, video processor outputs recorded video through output port 312, if desired, for display. The video input received through port 305 is also sent to digitizer 315 where it is converted into a standard digital format for processing. Each frame of the digitized image 20 or, alternately, each of a selected subset of the frames, is compared in visual analysis module 320 to similarly digitized

images of known facial expressions or movements of the viewer stored in video library files 325. For a more complete discussion of the analysis of facial gestures, see Antonio Colminarez, *Modeling the Dynamics of Facial Expressions*, _____

5 (submitted to the Computer Vision & Pattern Recognition Conference to be held in Hawaii from December 11-13, 2001), which is incorporated herein by reference. The comparison may be done on a pixel-by-pixel basis (using all or a portion of the pixels) or by using any known image analysis algorithm. Preferably, the reference video images stored in library files 315 are those of the viewer personally, but may also include a catalog of reference views from others, as well. The purpose of the comparison step is to determine as accurately as possible if a captured viewer expression or movement can be recognized as one associated with a known emotion or other human reaction. It is understood, of course, that any such evaluation is predictive or approximate rather than certain, and no implication is to be taken herein that a certain determination is claimed or even possible. Nevertheless, the information provided by video processor 300 provides important

10 input when evaluated in context. Matches detected by visual analysis module 320 are reported to sensor-signal receiver 285. If

15

20

no matches are found, the report may indicate this condition or simply be omitted.

FIGURE 4 is a system diagram illustrating the interaction of sensory inputs according to one embodiment of the present invention, as well as possible applications for using the gathered and processed sensory data. Again, this embodiment is illustrative and not limiting. As described above, facial expressions and head movements are analyzed against reference images to determine recognizable emotions (block 405). Indicators such as smiles and the intensity of a gaze (that is, duration without moving and the direction of gaze) are interpreted according to a set of predetermined criteria. Gross actions are also interpreted, such as movement around the room (block 410) as detected by video camera 240 or motion sensor 230. In addition, biometric response data captured by sensors on, for example, chair 215 or sensor shirt 220 (block 415) may be characterized as associated with certain human emotional responses in a manner similar to that described above in association with captured video images, albeit with somewhat less certainty in most cases. Note that certain sensory data, such as body temperature, may simply be analyzed by noting changes against a baseline or previous level. This type of analysis is certainly

easier than the video image comparisons, though in one preferred embodiment it also takes into account changes in environmental conditions, for example room temperature, outside weather conditions, ambient noise level, time of day, and the like.

5 Other inputs may, of course, also be taken into consideration, for example, the viewer's action in changing programming selections, especially where provided with a remote control, such as remote control 125 (shown in FIGURE 1), making such changes easy to execute (block 420). The programming changes may indicate a distaste for a particular program segment, or if frequent and without apparent direction, may indicate general boredom. In some cases, a viewer may engage in "channel surfing" where numerous channels are reviewed for short intervals in rapid succession. Even where part of an overall pattern, however, a longer than expected pause to view a particular segment may be informative. And the physical force being applied with each press of the "next-channel" button, as measured by a deflection gauge or similar device in the remote control 125, may indicate boredom or frustration - especially when coupled with the sensing and 10 recognition of certain contemporaneous audio responses, such as 15 sighs or expletives. As shall be apparent, experience and 20

observation while practicing the system and method of the present invention will enhance the ability to correlate certain sensory inputs, alone or in combination, with certain human emotional responses.

5 Each of the sensory-input features described above, and any others available as well, are fed directly or indirectly for classification of the particular behavior or condition involved (block 425). The individual inputs are each examined to determine if a discrete recognizable condition or change in condition can be segregated from the continuous, or at least intermittent stream of related input. Once a discrete sensory event is classified, it can be compared with a knowledge base of data in an attempt to correlate with a known mood, emotion, or other reaction (block 430). It may also be possible to determine how intense the particular mood or emotion is.

10 In a preferred embodiment, the extracted facial expressions from the visual domain are all labeled by symbols and given values for the feature vector. For example, if there are five states: sad, laughing, indifferent, bored, afraid, then these are symbolically encoded (preferably with numbers). Also, the volume of the emotion is recorded (e.g. from mild to extreme happiness) on a scale from 0

to 1, and the degree of certainty of the algorithm is also recorded. The viewer's pose and actions are tracked and these are also recorded: sitting, getting up, distracted reading a newspaper, shouting, etc. (these are also specific values for our feature vector). The Biometric responses are also recorded: electrocardiogram, electromyogram, respiration and skin conductance sensors that measure the autonomic nervous system activation. These signals are digitized in real time and recorded. All these features are not necessarily useful. First in the learning mode, a linear discriminate function can be used to rank each feature individually based on recognition performance to find an optimal set of features for recognizing patterns. Also, the remote control response pattern can fall into several categories and classified into: "channel zapping", active watching (with volume up, rewind, or slow mode), semi-active watching, non-active. All these values are used in the feature vector $f=(f_1, f_2, \dots, f_n)$, and the feature vector is recorded for a time segment (e.g. every 2 seconds).

Next, these observation symbols are fed into a Hidden Markov Model. Hidden Markov Model (HMM) is a common technique widely used in signal processing. The essence of HMM is to construct a model that explains the occurrence of observations (symbols) and use it

to identify other observation sequences. The fundamentals of HMM and its applications are presented in L. R. Rabiner, *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*, 77 PROCEEDINGS OF THE IEEE 257-285 (1989), which is
5 incorporated herein by reference.

Briefly stated, In an HMM, there are a finite number of states and the HMM is always in one of those states. At each clock time, it enters a new state based on a transition probability distribution depending on the previous state. After
10 a transition is made, an output symbol is generated based on a probability distribution depending on the current state. In the formal definition of HMM, the states are denoted as $Q=\{q_1, q_2, \dots, q_N\}$, where N is the number of states and the observation symbols are denoted as $V=\{v_1, v_2, \dots, v_M\}$, where M is the number of
15 observation symbols. The transition probability distribution between states is represented by a matrix $A=\{a_{ij}\}$, where $a_{ij}=\Pr\{q_j$ at $t+1 | q_i$ at $t\}$, and the observation symbol probability distribution is represented by the matrix $B=\{b_j(k)\}$, where $b_j(k)$ is the probability of generating observation v_k when the current
20 state is q_j .

The system consists of two phases, namely training and classification. We construct different HMM's, corresponding to different behavior (e.g. strong liking, indifference, disgust, appalled), through training with a collection of feature values as
5 explained above. The HMM training is essentially adjusting parameters of $\lambda=(A, B, \pi)$ to maximize the probability of the observation sequences $\text{Pr}(O|\lambda)$. Here π stands for the initial state distribution and is defined as $\pi=\{\pi_i\}$, where π_i is the probability of state q_i being the initial state of the HMM. O is the observation sequence.

In the classification phase, the observation sequence consisting of high level labels is extracted from a given set of input feature values for facial, biometric, person action, and remote control response patterns. Then the sequence is fed to the
15 different HMM's (e.g. for strong liking, indifference, disgust, appalled) as input and is classified as the class of the HMM that generates the highest response (probability of the observation).

Once the behavior of the viewer is identified, it is fed back into the system along with the type of the program and program
20 segment currently being viewed by the viewer. Inferences are made as to whether the user likes, dislikes or is neutral to the program

being shown. This is then used to present a different program (or some action to improve the programming content like augmentation etc.) to the viewer that might be better enjoyed based on the viewing profile of the user. The response to this new information 5 is monitored to learn the kind of information or action that the user prefers in different cases. In this way, a combination of the behavior is used for mood inference, type, intensity, and the like, in block 430.

Along with this viewer condition analysis, the program
10 input to which the viewer is being exposed is also analyzed
(block 435). Electronic program guide (EPG) information is
descriptive data relating to the remainder of programming input
available from a given source. While some of this information
is used to generate an actual program guide display that may be
15 used by a viewer in making a channel selection, other
information is used by the system of the present invention for
associating displayed program content with witnessed viewer
behavior. In other words, the EPG provides electronic cues to
the system indicating when a particular program has started or
20 finished. Note that the use of program segments results from a
recognition that even a single television program or motion

picture is not a homogeneous event, but rather a series of segments (or "scenes"). The process of video segmentation and selection at subprogram level is explained more fully in U.S.

Patent Application No. 09/442,960, entitled METHOD AND APPARATUS FOR

5 AUDIO/DATA/VISUAL INFORMATION SELECTION, filed by Nevenka Dimitrova,
Thomas McGee, Herman Elenbaas, Lalitha Agnihotri, Radu
Jasinschi, Serhan Dagtas, Aaron Mendelsohn on November 18, 1999,
co-owned by the Assignee of the present application, and
incorporated herein by reference. Further explanation is also
provided in R. S. Jasinschi, N. Dimitrova, T. McGee, L.
Agnihotri, J. Zimmerman, & D. Li, *Integrated Multimedia
Processing for Topic Segmentation and Classification*, PROCEEDINGS
OF THE IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (Thessaloniki,
October 2001), which is also incorporated herein by reference.

15 The viewer response to each of these segments may well be a
more useful estimator of viewer preferences than an evaluation
of the program as a whole. It is in this light that the system
and method of the present invention is most advantageously used.

In one particularly preferred embodiment, a programming
20 provider supplies (presumably for a fee) at least some sensory
equipment to a programming subscriber (that is, a viewer), or

provides instruction on how to connect electrical devices commonly found in subscribers' home (such as a video camera) to provide sensory readings. The sensory signals, processed or unprocessed, are returned to the provider, who analyzes them and uses them to
5 either adjust the programming content being set to the viewer or to create a channel selection control for the subscriber's (viewer's) use that assists the viewer in selecting programming similar to that which produced positive responses. In this situation, of course, the EPG information preferably includes program
10 segmentation information so that a sensed viewer reaction can be associated with a particular program segment.

Instead of, or in combination with, using EPG information to segment programming (i.e., divide it up logically into segments to be compared with discrete viewer responses), a segmenting function
15 may also be included, for example, in set top box 150, to automatically divide the program appropriately for analysis. The program (and program segmenting) information is compared to the moods and emotions determined to occur during the program segments, providing important program preference information (block 440).
20 This preference information can be used in a variety of ways. First, of course, it can be stored in a memory recorder (block

445), and reviewed later as another aid to predicting future viewer preferences. This may be done in a test situation or, as previously described, where the system operates in a viewer's home, the information may be transmitted to the programming provider for 5 individual or aggregate (that is, test-audience) analysis. As regards the individual viewer, the learned preferences information may also be processed (block 450) for use in future mood or emotion determinations (block 430). In addition, the preference information may be used, based on the assumption that the user 10 would choose programming content that would appear similar to previously-viewed content that produced a favorable reaction, if the viewer is given the opportunity to do so. This effect may be accomplished by a simple alert, notifying the viewer that a particular type of programming is now available (block 455). Or a 15 number of similar upcoming programs may be offered as a recommendation (block 460). In either case, of course, the user is simply offered the opportunity to view or record the recommended programming. In an alternate embodiment, however, the system may simply select certain programming for recording, automatically 20 sending it to a recording device (block 465). This is especially useful where there is the capability to record a great deal of

material, so that the programming may be saved for some time and not simply overwritten in the next recording session. Finally, the actual content of the viewed or recorded programming may be augmented based on previously learned user preferences (block 470).

5 The augmenting material is presumably available from the programming provider, but used according to the individual tastes of the viewer. For example, a viewer who has previously demonstrated a preference for action scenes may have a car chase inserted in place of a courtroom scene, while another viewer who prefers comedy may have a humorous sequence instead. If the viewers are willing to indicate their identity in some fashion, of course, or if identity can be automatically determined by system sensors, the system may customize programming to their individual tastes (or to composite tastes, for multiple viewers that frequently watch together).

10 FIGURE 5 is a flowchart illustrating a method 500 of collecting, processing, and using sensory-feedback information according to an embodiment of the present invention. Initially (process step START), it is assumed that the programming reception and display system and the sensory feedback system is in place. As 15 the program display begins, sensory devices are powered up and

activated (process step 505). The system immediately receives initial input and, in the illustrated embodiment, establishes the initial input as a baseline (process step 510). In an alternate embodiment, a baseline for each identifiable viewer is already established, and a new baseline is determined only where the identity of a viewer cannot be determined. Sensor input is processed until a recognizable sensory condition is detected (such as a rise in temperature or perspiration level) (process step 515), at which time the indication is transmitted to the sensory processor for mood/emotion determination is associated with information regarding the program segment that was being displayed when the sensory result of the mood/emotion was detected (process step 520). If the response to the programming was within an expected range, the preference information is stored in a database (process step 525). Note that the preference information is not "within expected range" if it is wholly spurious, such as a normally somber viewer laughing through an unhappy segment - in which case some distraction is suspected. In a preferred embodiment, such responses are disregarded. The process may also continue with the steps of sending the preference information back to the provider (process step 530) or of using the preference

information locally (process step 535) or both. (Several examples of local use have been provided above in reference to FIGURE 4.)

Although the present invention has been described in detail, those skilled in the art should understand that they can make 5 various changes, substitutions and alterations herein without departing from the spirit and scope of the invention in its broadest form.